

Performance evaluation for ML sequence detection in ISI channels with Gauss Markov Noise

Naveen Kumar*, Aditya Ramamoorthy* and Murti Salapaka†

*Dept. of Electrical and Computer Engineering

Iowa State University, Ames, IA 50010

Email: nk3,adityar@iastate.edu

†Dept. of Electrical and Computer Engineering

University of Minnesota, Minneapolis, MN 55455

Email: murtis@umn.edu

Abstract—Inter-symbol interference (ISI) channels with data dependent Gauss Markov noise have been used to model read channels in magnetic recording and other data storage systems. The Viterbi algorithm can be adapted for performing maximum likelihood sequence detection in such channels. However, the problem of finding an analytical upper bound on the bit error rate of the Viterbi detector in this case has not been fully investigated. Current techniques rely on an exhaustive enumeration of short error events and determine the BER using a union bound.

In this work, we consider a subset of the class of ISI channels with data dependent Gauss-Markov noise. We derive an upper bound on the pairwise error probability (PEP) between the transmitted bit sequence and the decoded bit sequence that can be expressed as a product of functions depending on current and previous states in the (incorrect) decoded sequence and the (correct) transmitted sequence. In general, the PEP is asymmetric. The average BER over all possible bit sequences is then determined using a pairwise state diagram. Simulations results which corroborate the analysis of upper bound, demonstrate that analytic bound on BER is tight in high SNR regime. In the high SNR regime, our proposed upper bound obviates the need for computationally expensive simulation.

I. INTRODUCTION

Maximum likelihood sequence detection (MLSD) in channels with inter-symbol-interference and data dependent time-correlated noise is an important problem in many domains. For example, in magnetic recording, the statistics of percolation and nonlinear effects between transitions [11], [6] result in noise that exhibits data-dependent time-correlation. Recently, similar noise models for nanotechnology based probe storage have also been developed and the corresponding detectors have been found to have significantly improved performance compared to the current state of the art [7]. It is well-recognized that a sequence detector designed for an AWGN ISI model can have a significant loss of performance if the data dependence and time-correlation of the noise is not taken into account.

In the case of finite ISI channels with memoryless noise, Forney [3], [4] presented an MLSD solution based on the Viterbi algorithm. Upper bounds on the error probability of the detector can be derived based on flowgraph techniques [2], [8], [10]. The work of Kavcic & Moura [1] considered channels with finite ISI and noise modeled by a finite memory Gauss-Markov process. The work of [1], also presents certain

approaches (see section V in [1]) for computing an upper bound on the performance of the detector. However, their technique is not based on flowgraph techniques, and requires an enumeration of all error events of relevant lengths and an estimate of the corresponding pairwise error probability upper bound. We emphasize that an analytical technique for estimating detector performance is of great value since it allows us to predict the performance at high SNR's where simulation can be time-consuming.

In an ISI channel with additive white Gaussian noise (AWGN), the pairwise error probability (PEP) between two state sequences can be easily factorized as a product of functions depending on current and previous states in the (incorrect) decoded sequence and the (correct) transmitted sequence. Let \bar{S} and \hat{S} be the transmitted and decoded state sequences respectively. Then this means that the probability that the detector prefers \hat{S} to \bar{S} , is denoted by $P(\hat{S}|\bar{S}) = \prod_{k=0}^{N-1} h(\hat{S}_{k-1}^k, S_{k-1}^k)$ where h is a function of current state and previous decoded states $\hat{S}_{k-1}^k = (\hat{S}_{k-1}, \hat{S}_k)$ and actual states $S_{k-1}^k = (S_{k-1}, S_k)$. Moreover, the PEP is symmetric due to the symmetric nature of white Gaussian noise, i.e., $P(\hat{S}|\bar{S}) = P(\bar{S}|\hat{S})$. Together, these properties allow the application of the error state diagram method for finding an upper bound on the BER [10].

In contrast, for the ISI channel with data-dependent Gauss-Markov noise (considered in [1]), neither of these properties hold. The signal dependent and time-correlated noise makes the PEP asymmetric. Further the PEP does not factorize in a suitable manner as required for the application of flowgraph techniques. This makes the estimation of BER for such channels, quite challenging.

Main Contributions: In this paper, we consider a subset of the class of ISI channels with Gauss-Markov noise. For these channels, we arrive at an upper bound to the PEP that can be expressed as a product of functions depending on current and previous states in the (incorrect) decoded sequence and the (correct) transmitted sequence. The asymmetric character of the PEP, i.e., the fact that $P(\hat{S}|\bar{S}) \neq P(\bar{S}|\hat{S})$ necessitates an average over all correct and erroneous state sequences. We show that this can be achieved using the concept of the “pairwise state diagram” [2]. Based on this, we present an

analytical technique for determining an upper bound on the BER. Simulations results show that our proposed bound is tight in the high SNR regime.

The paper is organized as follows. Section II introduces the channel model and describes the corresponding Viterbi decoding algorithm. Section III presents an upper bound on the detector BER. Section IV demonstrates simulation results that confirm the analytical bounds. Section V summarizes the main findings of this paper and outlines future work.

II. CHANNEL MODEL AND VITERBI DETECTOR

We introduce the channel model and the corresponding detector in this section. A word about notation. In what follows, if z_k is a discrete-time indexed sequence at k^{th} time instant, the column vector of sequence samples from time instant k_1 to k_2 is denoted by $z_{k_1}^{k_2} = [z_{k_1} \dots z_{k_2}]^T$ where $k_1 \leq k_2$. We will use the notation $f(\cdot|\cdot)$ to denote a conditional pdf. The precise pdf under consideration will be evident from the context of the discussion.

A. Channel Model

Let a_k denote the k^{th} source bit that is equally likely to be 0 or 1. The channel output shown in Figure 1 with intersymbol interference (ISI) of length I is given by,

$$z_k = y(a_{k-I}^k) + n_k,$$

where $y(a_{k-I}^k)$ is the noiseless channel output dependent only on the $I+1$ past transmitted bits. The noise n_k is modeled as a signal dependent Gauss-Markov noise process with memory length L as explained below.

$$n_k = \bar{b}^T n_{k-L}^{k-1} + \sigma(a_{k-I}^k) w_k,$$

where the vector \bar{b} represents L coefficients of an autoregressive filter, $\sigma(a_{k-I}^k)$ is signal dependent parameters and w_k is a zero mean unit variance i.i.d Gaussian random variable. Note that in the most general model (considered in [1]), even the autoregressive filter \bar{b} would depend on the data sequence \bar{a} . However, in this work, we only work with models where \bar{b} is fixed. We revisit this point in Section III. The noise n_k can be rewritten as,

$$\begin{aligned} n_k &= \bar{b}^T \begin{pmatrix} n_{k-L} \\ \vdots \\ n_{k-1} \end{pmatrix} + \sigma(a_{k-I}^k) w_k \\ &= \bar{b}^T \begin{pmatrix} z_{k-L} - y(a_{k-L-I}^{k-1}) \\ \vdots \\ z_{k-1} - y(a_{k-1-I}^{k-1}) \end{pmatrix} + \sigma(a_{k-I}^k) w_k \end{aligned}$$

This implies that

$$z_k = y(a_{k-I}^k) + \bar{b}^T \begin{pmatrix} z_{k-L} - y(a_{k-L-I}^{k-1}) \\ \vdots \\ z_{k-1} - y(a_{k-1-I}^{k-1}) \end{pmatrix} + \sigma(a_{k-I}^k) w_k$$

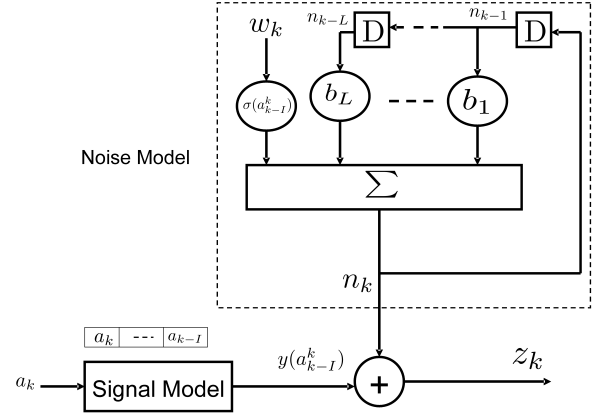


Fig. 1. Channel model with Gauss-Markov noise.

From above analysis, we can conclude that

$$f(z_k | z_0^{k-1}, \bar{a}) = f(z_k | z_{k-L}^{k-1}, a_{k-L-I}^k), \quad (1)$$

where we recall that $f(\cdot|\cdot)$ represents the conditional pdf.

B. Viterbi Detector

The maximum likelihood estimate of the bit sequence denoted \hat{a} is given by

$$\begin{aligned} \hat{a} &= \arg \max_{\bar{a} \in \{0,1\}^N} f(\bar{z}|\bar{a}) \\ &= \arg \max_{\bar{a} \in \{0,1\}^N} \prod_{k=0}^{N-1} f(z_k | z_0^{k-1}, \bar{a}) \\ &= \arg \max_{\bar{a} \in \{0,1\}^N} \prod_{k=0}^{N-1} f(z_k | z_{k-L}^{k-1}, a_{k-L-I}^k) \quad (\text{Using (1)}) \\ &= \arg \max_{\bar{a} \in \{0,1\}^N} \prod_{k=0}^{N-1} \frac{f(z_{k-L}^k | a_{k-L-I}^k)}{f(z_{k-L}^{k-1} | a_{k-L-I}^k)} \end{aligned}$$

We define a state $S_k = a_{k-L-I+1}^k$ (there will be a total of 2^{L+I} states). With this definition, $f(z_{k-L}^k | a_{k-L-I}^k) = f(z_{k-L}^k | S_{k-1}^k)$. Moreover it is Gaussian distributed,

$$f(z_{k-L}^k | S_{k-1}^k) \sim N(\bar{\mu}(S_{k-1}^k), C(S_{k-1}^k))$$

where $\bar{\mu}(S_{k-1}^k)$ is the mean vector and $C(S_{k-1}^k)$ is the covariance matrix.

With our state definition, we can reformulate the detection problem as the following MLSD problem.

$$\begin{aligned} \hat{\hat{S}} &= \arg \max_{\hat{S}} \prod_{k=0}^{N-1} \frac{f(z_{k-L}^k | a_{k-L-I}^k)}{f(z_{k-L}^{k-1} | a_{k-L-I}^k)} \\ &= \arg \max_{\hat{S}} \prod_{k=0}^{N-1} \frac{f(z_{k-L}^k | S_{k-1}^k)}{f(z_{k-L}^{k-1} | S_{k-1}^k)} \\ &= \arg \min_{\hat{S}} \sum_{k=0}^{N-1} \log \frac{|C(S_{k-1}^k)|}{|c(S_{k-1}^k)|} \\ &\quad + (z_{k-L}^k - \bar{\mu}(S_{k-1}^k))^T C(S_{k-1}^k)^{-1} (z_{k-L}^k - \bar{\mu}(S_{k-1}^k)) \\ &\quad - (z_{k-L}^{k-1} - \bar{\mu}'(S_{k-1}^k))^T c(S_{k-1}^k)^{-1} (z_{k-L}^{k-1} - \bar{\mu}'(S_{k-1}^k)) \end{aligned}$$

where $\hat{\hat{S}}$ is the estimated state sequence, $c(S_{k-1}^k)$ is the upper $L \times L$ principal minor of $C(S_{k-1}^k)$ and $\bar{\mu}'(S_{k-1}^k)$ collects the

first L elements of $\bar{\mu}(S_{k-1}^k)$. It is assumed that the first state is known. With the metric given above, Viterbi decoding can be applied to get the ML state sequence [9] and the corresponding bit sequence.

The matrix $C(S_{k-1}^k)$ is of dimension $(L+1) \times (L+1)$. For higher values of L , the complexity of detector increases as the decoding metric involves the inversion of the matrix $C(S_{k-1}^k)$. However, the matrix inversion lemma can be used here to obtain

$$C(S_{k-1}^k)^{-1} = \begin{bmatrix} \mathbf{c}(S_{k-1}^k) & \bar{c} \\ \bar{c}^T & c \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{c}(S_{k-1}^k)^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \frac{\bar{w}(S_{k-1}^k)\bar{w}(S_{k-1}^k)^T}{\gamma(S_{k-1}^k)}, \quad (2)$$

where

$$\bar{w}(S_{k-1}^k) = \begin{bmatrix} -\mathbf{c}(S_{k-1}^k)^{-1}\bar{c} \\ 1 \end{bmatrix} = \begin{bmatrix} -\bar{b} \\ 1 \end{bmatrix}, \text{ and}$$

$$\gamma(S_{k-1}^k) = (c - \bar{c}^T \mathbf{c}(S_{k-1}^k)^{-1} \bar{c}) = \sigma^2(a_{k-I}^k).$$

Using (2), we can simplify the detector as follows.

$$\hat{S} = \arg \min_{\bar{S}} \sum_{k=0}^{N-1} \log \sigma^2(a_{k-I}^k) + \frac{([- \bar{b}^T \ 1](z_{k-L}^k - \bar{\mu}(S_{k-1}^k)))^2}{\sigma^2(a_{k-I}^k)}.$$

It should be noted that the above expression does not involve any matrix inversion. This reduces the complexity of the detector substantially. Another observation is that the Viterbi decoding metric involves passing z_{k-L}^k through a filter $[- \bar{b}^T \ 1]$ which is the inverse of the autoregressive filter of noise process n_k shown in Figure 1. The metric first uncorrelates the noise with an FIR filter and then applies the Euclidean metric to the output of the filter.

III. UPPER BOUND ON BER

As discussed previously, the channel model under consideration (cf. Section II), is such that the corresponding PEP is asymmetric, and moreover does not factorize as a product of appropriate functions as required by flowgraph techniques. We now show that we can address this issue by using the Gallager upper bounding technique [5], coupled with a suitable change of variables.

Denote an error event of length N as $\epsilon_N = (\bar{S}, \hat{S})$ such that \bar{S} and \hat{S} are valid state sequences and $S_k = \hat{S}_k$, $S_{k+N} = \hat{S}_{k+N}$, $S_{k+j} \neq \hat{S}_{k+j}$ for $1 \leq j \leq N-1$ and $S_{k+j} = \hat{S}_{k+j}$ for other values of j where \hat{S}_k and S_k are the estimated and correct state respectively. Using this, an upper bound on the BER can be found as follows [2],

$$P_b(e) \leq \sum_{N=1}^{\infty} \sum_{\bar{S}} P(\bar{S}) \sum_{\hat{S}: (\bar{S}, \hat{S}) \in E_N} \nu(\bar{S}, \hat{S}) P(\hat{S}|\bar{S}),$$

where $\nu(\bar{S}, \hat{S})$ is the number of erroneous bits along the sequences \bar{S} and \hat{S} and E_N is the set of all error events ϵ_N of length N . The number of erroneous bits is given by

$$\nu(\bar{S}, \hat{S}) = \frac{d}{dZ} [\Pi_{k=0}^{N-1} Z^{\delta(a_k, \hat{a}_k)}] \Big|_{Z=1}$$

where $\delta(a_k, \hat{a}_k) = 1$, if $a_k \neq \hat{a}_k$ and Z is a dummy variable. Using this the upper bound above can be expressed as

$$P_b(e) \leq \sum_{N=1}^{\infty} \sum_{\bar{S}} P(\bar{S}) \sum_{\hat{S}: (\bar{S}, \hat{S}) \in E_N} \frac{d}{dZ} \cdot [\Pi_{k=0}^{N-1} Z^{\delta(a_k, \hat{a}_k)}] \Big|_{Z=1} P(\hat{S}|\bar{S})$$

where $P(\bar{S}) = P(S_0)P(S_1|S_0) \dots P(S_N|S_{N-1}) = \frac{1}{M} \cdot \frac{1}{2^N}$ if \bar{S} is valid state sequence, (M is the number of states). The upper bound on the PEP can be using Gallager's technique [5] as shown below. Let $A(\bar{S}, \hat{S}) = \{\bar{z} : f(\bar{z}|\hat{S}) \geq f(\bar{z}|\bar{S})\}$. Note that using previous arguments, we also have that $A(\bar{S}, \hat{S}) = \{\bar{z} : \Pi_{k=0}^{N-1} \frac{f(z_k|\hat{S}_{k-1}^k, z_{k-L}^{k-1})}{f(z_k|S_{k-1}^k, z_{k-L}^{k-1})} \geq 1\}$. Now,

$$\begin{aligned} P(\hat{S}|\bar{S}) &= P(\hat{S}_0 \dots \hat{S}_{N-1} | S_0 \dots S_{N-1}) \\ &= \int_{A(\bar{S}, \hat{S})} \Pi_{k=0}^{N-1} f(z_k | S_{k-1}^k, z_{k-L}^{k-1}) d\bar{z} \\ &\leq \min_{\forall \rho_k} \int \Pi_{k=0}^{N-1} f(z_k | S_{k-1}^k, z_{k-L}^{k-1}) \\ &\quad \cdot \Pi_{k=0}^{N-1} \left(\frac{f(z_k | \hat{S}_{k-1}^k, z_{k-L}^{k-1})}{f(z_k | S_{k-1}^k, z_{k-L}^{k-1})} \right)^{\rho_k} d\bar{z} \\ &= \min_{\forall \rho_k} \int \Pi_{k=0}^{N-1} (f(z_k | S_{k-1}^k, z_{k-L}^{k-1}))^{1-\rho_k} \\ &\quad \cdot (f(z_k | \hat{S}_{k-1}^k, z_{k-L}^{k-1}))^{\rho_k} d\bar{z} \end{aligned}$$

where $0 \leq \rho_k \leq 1$ for $k = 0, \dots, N-1$.

The above integral can be simplified as follows.

$$\begin{aligned} &\int \Pi_{k=0}^{N-1} (f(z_k | S_{k-1}^k, z_{k-L}^{k-1}))^{1-\rho_k} \\ &\quad \cdot (f(z_k | \hat{S}_{k-1}^k, z_{k-L}^{k-1}))^{\rho_k} d\bar{z} \\ &= \int \Pi_{k=0}^{N-1} \frac{1}{\sqrt{2\pi}\sigma^{1-\rho_k}(a_{k-I}^k)^{\hat{\sigma}\rho_k}(\hat{a}_{k-I}^k)} \\ &\quad \cdot \exp\left(-\frac{(1-\rho_k)([- \bar{b}^T \ 1](z_{k-L}^k - \bar{\mu}(S_{k-1}^k)))^2}{2\sigma^2(a_{k-I}^k)}\right. \\ &\quad \left.- \frac{\rho_k([- \bar{b}^T \ 1](z_{k-L}^k - \hat{\mu}(\hat{S}_{k-1}^k)))^2}{2\hat{\sigma}^2(\hat{a}_{k-I}^k)}\right) d\bar{z} \\ &= \Pi_{k=0}^{N-1} \int \frac{1}{\sqrt{2\pi}\sigma^{1-\rho_k}(a_{k-I}^k)^{\hat{\sigma}\rho_k}(\hat{a}_{k-I}^k)} \\ &\quad \cdot \exp\left(-\frac{(1-\rho_k)(u_k - \mathfrak{M}(S_{k-1}^k))^2}{2\sigma^2(a_{k-I}^k)}\right. \\ &\quad \left.- \frac{\rho_k(u_k - \hat{\mathfrak{M}}(\hat{S}_{k-1}^k))^2}{2\hat{\sigma}^2(\hat{a}_{k-I}^k)}\right) du_k \end{aligned}$$

where $u_k = [- \bar{b}^T \ 1] \cdot [z_{k-L} \dots z_k]^T$, $\mathfrak{M}(S_{k-1}^k) = [- \bar{b}^T \ 1] \cdot \bar{\mu}(S_{k-1}^k)$, $\hat{\mathfrak{M}}(\hat{S}_{k-1}^k) = [- \bar{b}^T \ 1] \cdot \hat{\mu}(\hat{S}_{k-1}^k)$. The Jacobian matrix for the change of variables has determinant equal to 1, since the corresponding matrix of partial derivatives has ones on

the diagonal and is lower triangular. Note that the change of variables decouples the original expression, so that it can be expressed as the product of N independent integrals. Now we can simplify the PEP as follows.

$$\begin{aligned}
P(\hat{a}|\bar{a}) &\leq \min_{\forall \rho_k} \Pi_{k=0}^{N-1} \int \frac{1}{\sqrt{2\pi}\sigma^{1-\rho_k}(a_{k-I}^k)\hat{\sigma}^{\rho_k}(\hat{a}_{k-I}^k)} \\
&\cdot \exp\left(-\frac{(1-\rho_k)(u_k - \mathfrak{M}(S_{k-1}^k))^2}{2\sigma^2(a_{k-I}^k)}\right. \\
&\quad \left.- \frac{\rho_k(u_k - \hat{\mathfrak{M}}(\hat{S}_{k-1}^k))^2}{2\hat{\sigma}^2(\hat{a}_{k-I}^k)}\right) du_k \\
&= \Pi_{k=0}^{N-1} \min_{\rho_k} \int \frac{1}{\sqrt{2\pi}\sigma^{1-\rho_k}(a_{k-I}^k)\hat{\sigma}^{\rho_k}(\hat{a}_{k-I}^k)} \\
&\cdot \exp\left(-\frac{(1-\rho_k)(u_k - \mathfrak{M}(S_{k-1}^k))^2}{2\sigma^2(a_{k-I}^k)}\right. \\
&\quad \left.- \frac{\rho_k(u_k - \hat{\mathfrak{M}}(\hat{S}_{k-1}^k))^2}{2\hat{\sigma}^2(\hat{a}_{k-I}^k)}\right) du_k \quad (3) \\
&= \Pi_{k=0}^{N-1} \min_{\rho_k} \frac{\sigma^{\rho_k}(a_{k-I}^k)\hat{\sigma}^{1-\rho_k}(\hat{a}_{k-I}^k)}{\sqrt{(1-\rho_k)\hat{\sigma}^2(\hat{a}_{k-I}^k) + \rho_k\sigma^2(a_{k-I}^k)}} \\
&\cdot \exp\left(-\frac{(1-\rho_k)\hat{\sigma}^2(\hat{a}_{k-I}^k)\mathfrak{M}^2(S_{k-1}^k) + \rho_k\sigma^2(a_{k-I}^k)\hat{\mathfrak{M}}^2(\hat{S}_{k-1}^k)}{2\sigma^2(a_{k-I}^k)\hat{\sigma}^2(\hat{a}_{k-I}^k)}\right. \\
&\quad \left.+ \frac{((1-\rho_k)\hat{\sigma}^2(\hat{a}_{k-I}^k)\mathfrak{M}(S_{k-1}^k) + \rho_k\sigma^2(a_{k-I}^k)\hat{\mathfrak{M}}(\hat{S}_{k-1}^k))^2}{2\sigma^2(a_{k-I}^k)\hat{\sigma}^2(\hat{a}_{k-I}^k)((1-\rho_k)\hat{\sigma}^2(\hat{a}_{k-I}^k) + \rho_k\sigma^2(a_{k-I}^k))}\right) \\
&= \Pi_{k=0}^{N-1} W(S_{k-1}^k, \hat{S}_{k-1}^k) \quad (4)
\end{aligned}$$

where $W(S_{k-1}^k, \hat{S}_{k-1}^k)$ is a function of $\sigma(a_{k-I}^k)$, $\hat{\sigma}(\hat{a}_{k-I}^k)$, $\mathfrak{M}(S_{k-1}^k)$ and $\hat{\mathfrak{M}}(\hat{S}_{k-1}^k)$ and the simplification of the integral in (3) is given in the Appendix.

It is important to note that the factorization of PEP given by (4) for our channel model is possible because the autoregressive filter \bar{b} is not dependent on the input bit sequence. In [1], \bar{b} is assumed to be data dependent given by $\bar{b}(a_{k-I}^k)$. When the autoregressive filter $\bar{b}(a_{k-I}^k)$ becomes data dependent, it is very difficult to write PEP in the form given in (4). In this case, the inverse of the autoregressive filter of the noise process n_k ($[-\bar{b}(a_{k-I}^k)^T \ 1]$) is state-dependent which means that the actual state transition (S_{k-1}^k) and estimated state transition (\hat{S}_{k-1}^k) have different filters. In this situation, the specific change of variables used above does not seem to work.

Probability of bit error can now be further simplified as [2],

$$\begin{aligned}
P_b(e) &\leq \sum_{N=1}^{\infty} \sum_{\bar{S}} P(\bar{S}) \sum_{\hat{S}: (\bar{S}, \hat{S}) \in E_N} \frac{d}{dZ} \\
&\cdot [\Pi_{k=0}^{N-1} Z^{\delta(a_k, \hat{a}_k)}|_{Z=1} P(\hat{S}|\bar{S})] \\
&= \frac{d}{dZ} \sum_{N=1}^{\infty} \frac{1}{M} \cdot \frac{1}{2^N} \sum_{\bar{S}} \sum_{\hat{S}} \Pi_{k=0}^{N-1} Z^{\delta(a_k, \hat{a}_k)} P(\hat{S}|\bar{S})|_{Z=1}
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{d}{dZ} \sum_{N=1}^{\infty} \frac{1}{M} \cdot \frac{1}{2^N} \sum_{\bar{S}} \sum_{\hat{S}} \Pi_{k=0}^{N-1} Z^{\delta(a_k, \hat{a}_k)} \\
&\cdot W(S_{k-1}^k, \hat{S}_{k-1}^k)|_{Z=1} \quad (\text{Using (4)}) \\
&= \frac{1}{M} \frac{d}{dZ} \sum_{N=1}^{\infty} \sum_{\bar{S}} \sum_{\hat{S}} \Pi_{k=0}^{N-1} \frac{1}{2} Z^{\delta(a_k, \hat{a}_k)} W(S_{k-1}^k, \hat{S}_{k-1}^k)|_{Z=1} \\
&= \frac{1}{M} \frac{d}{dZ} (T(Z))|_{Z=1}
\end{aligned}$$

For obtaining $T(Z)$, we construct a product trellis. Consider a matrix $V(Z)$ of order $M^2 \times M^2$, where each row and column is indexed by a pair of states corresponding to the actual and the errored states. Let β_i represent a state that takes one of 2^{L+I} values. Consider the entry of $V(Z)$ indexed by $((\beta_i, \beta_j), (\beta'_i, \beta'_j))$,

$$\begin{aligned}
[V(Z)]_{((\beta_i, \beta_j), (\beta'_i, \beta'_j))} &= \begin{cases} \frac{1}{2} Z^{\delta(a_i, a_j)} W((\beta_i, \beta_j), (\beta'_i, \beta'_j)) \\ 0, \text{ if either } \beta_i \rightarrow \beta'_i \text{ or } \beta_j \rightarrow \beta'_j \text{ not allowed} \end{cases}
\end{aligned}$$

where $W((\beta_i, \beta_j), (\beta'_i, \beta'_j))$ can be found for state transitions $\beta_i \rightarrow \beta'_i$ and $\beta_j \rightarrow \beta'_j$ using (4) and a_i and a_j are latest bit in states β'_i and β'_j respectively. A product state is called good state if $\beta'_i = \beta'_j$ and bad otherwise. $V(Z)$ will have a structure which has $V_{GG}(Z)$ (good to good state transition), $V_{GB}(Z)$ (good to bad state transition), $V_{BG}(Z)$ (bad to good state transition) and $V_{BB}(Z)$ (bad to bad state transition),

$$V(Z) = \begin{bmatrix} V_{GG}(Z) & V_{GB}(Z) \\ V_{BG}(Z) & V_{BB}(Z) \end{bmatrix}$$

where the order of $V_{GG}(Z)$ matrix is $M \times M$ and the order of $V_{BB}(Z)$ is $(M^2 - M) \times (M^2 - M)$. Now we can write $T(Z)$ as,

$$T(Z) = \mathbf{a}(Z) + \mathbf{b}(Z)(I - V_{BB}(Z))^{-1} \mathbf{c}(Z)$$

where $\mathbf{a}(Z) = \mathbf{1}^T V_{GG}(Z) \mathbf{1}$, $\mathbf{b}(Z) = \mathbf{1}^T V_{GB}(Z)$ and $\mathbf{c}(Z) = V_{BG}(Z) \mathbf{1}$. The symbol $\mathbf{1}$ denotes a vector all of whose entries are 1 and I is identity matrix of order $(M^2 - M) \times (M^2 - M)$. Using the above result, we can compute $P_b(e)$ as [2],

$$\begin{aligned}
P_b(e) &\leq \frac{1}{M} [\mathbf{a}'(1) + \mathbf{b}'^T(1)(I - V_{BB}(1))^{-1} \mathbf{c}(1) + \mathbf{b}^T(1) \\
&\cdot (I - V_{BB}(1))^{-1} \mathbf{c}'(1) + \mathbf{b}^T(1)(I - V_{BB}(1))^{-1} \\
&\cdot V'_{BB}(1)(I - V_{BB}(1))^{-1} \mathbf{c}(1)].
\end{aligned}$$

For our model, $V_{GG}(Z)$ is not a function of Z which means that $\mathbf{a}'(1) = 0$. Similarly, $\mathbf{c}(Z)$ is also not a function of Z which implies $\mathbf{c}'(1) = 0$ and it should also be noted that $\mathbf{b}'^T(1) = \mathbf{b}^T(1)$. The new bound for our channel model is,

$$\begin{aligned}
P_b(e) &\leq \frac{1}{M} [\mathbf{b}^T(1)(I - V_{BB}(1))^{-1} \mathbf{c}(1) + \mathbf{b}^T(1) \\
&\cdot (I - V_{BB}(1))^{-1} V'_{BB}(1)(I - V_{BB}(1))^{-1} \mathbf{c}(1)].
\end{aligned}$$

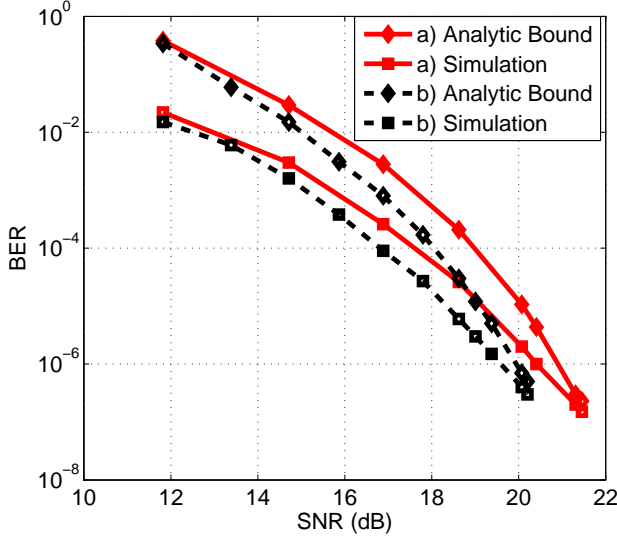


Fig. 2. BER with different SNR for the channel model with a) 8 states in decoding and b) 16 states in decoding.

IV. SIMULATION RESULTS

In the first set of simulations, we used the following parameters: $L = 2$ with $\bar{b} = [1 \ .5]$ and ISI memory $I = 1$. The signal dependent noise variance for 4 states are given by $\sigma^2(00) = 1$, $\sigma^2(01) = 2$, $\sigma^2(10) = 3$ and $\sigma^2(11) = 4$. The number of states in decoding is equal to 8 in this case. The SNR is defined as signal energy in $y(a_{k-I}^k)$ divided by total noise variance. We have used a linear signal component given as $y(a_{k-1}^k) = c(2a_k + a_{k-1})$ where the value of c can be varied to change the SNR. In Figure 2, the analytic bound follows the simulation BER. At an SNR of 21 dB, the analytic bound gives a BER equal to 3×10^{-7} whereas simulation BER is equal to 2×10^{-7} . The analytic bound is quite tight in high SNR regime. In another simulation, we used following parameters, $L = 3$ coefficients of an autoregressive filter is given by $\bar{b} = [1 \ .3 \ .5]$, ISI memory (I) is equal to 1 and signal dependent noise variance for 4 states are given by $\sigma^2(00) = 1$, $\sigma^2(01) = 2$, $\sigma^2(10) = 3$ and $\sigma^2(11) = 4$. The number of states in decoding is equal to 16 in this case. In Figure 2, the analytic bound again follows the simulation BER for this channel model with modified channel parameters. At an SNR of 20 dB, the analytic bound gives a BER equal to 7×10^{-7} whereas simulation BER is equal to 4×10^{-7} .

V. CONCLUSIONS AND FUTURE WORK

We considered the problem of deriving an analytical upper bound for ML sequence detection in ISI channels with signal dependent Gauss-Markov noise. In these channels the pairwise error probability (PEP) is not symmetric. Moreover, it is hard to express the PEP as a product of appropriate terms that allow the application of flowgraph techniques. In this work, we considered a subset of these channels, and demonstrated an appropriate upper bound on the PEP. Using this upper bound along with pairwise state diagrams, we arrive at analytical BER bounds that are tight in the high SNR regime. These bounds have been verified by our simulation results.

It would be interesting to examine whether our current techniques can be extended to address the general channel model. Moreover, it may be possible to reduce the complexity of evaluating the bound by reducing the size of the product trellis by exploiting channel characteristics. We are currently investigating these issues.

APPENDIX

The integral in the equation (3) can be expressed in the following form,

$$\begin{aligned} & \int \frac{1}{\sqrt{2\pi}\gamma} \exp\left(-\frac{1}{2}(\alpha(x-m)^2 + \beta(x-\hat{m})^2)\right) dx \\ &= \frac{1}{\gamma\sqrt{\alpha+\beta}} \exp\left(-\frac{\alpha m^2 + \beta \hat{m}^2}{2} + \frac{(\alpha m + \beta \hat{m})^2}{2(\alpha+\beta)}\right) \\ & \cdot \underbrace{\int \frac{\sqrt{\alpha+\beta}}{\sqrt{2\pi}} \exp\left(-\frac{\alpha+\beta}{2}\left(x - \frac{\alpha m + \beta \hat{m}}{\alpha+\beta}\right)^2\right) dx}_{=1} \\ &= \frac{1}{\gamma\sqrt{\alpha+\beta}} \exp\left(-\frac{\alpha m^2 + \beta \hat{m}^2}{2} + \frac{(\alpha m + \beta \hat{m})^2}{2(\alpha+\beta)}\right) \end{aligned}$$

where $\alpha = \frac{(1-\rho_k)}{\sigma^2(a_{k-I}^k)}$, $\beta = \frac{\rho_k}{\hat{\sigma}^2(\hat{a}_{k-I}^k)}$, $m = \mathfrak{M}(S_{k-1}^k)$, $\gamma = \sigma^{1-\rho_k}(a_{k-I}^k) \hat{\sigma}^{\rho_k}(\hat{a}_{k-I}^k)$ and $\hat{m} = \hat{\mathfrak{M}}(\hat{S}_{k-1}^k)$. Using the above equality, we can easily simplify the RHS of equation (3) as,

$$\begin{aligned} & \frac{1}{\sigma^{1-\rho_k}(a_{k-I}^k) \hat{\sigma}^{\rho_k}(\hat{a}_{k-I}^k)} \cdot \frac{\sigma(a_{k-I}^k) \hat{\sigma}(\hat{a}_{k-I}^k)}{\sqrt{(1-\rho_k) \hat{\sigma}^2(\hat{a}_{k-I}^k) + \rho_k \sigma^2(a_{k-I}^k)}} \\ & \cdot \exp\left(-\frac{(1-\rho_k) \hat{\sigma}^2(\hat{a}_{k-I}^k) \mathfrak{M}^2(S_{k-1}^k) + \rho_k \sigma^2(a_{k-I}^k) \hat{\mathfrak{M}}^2(\hat{S}_{k-1}^k)}{2\sigma^2(a_{k-I}^k) \hat{\sigma}^2(\hat{a}_{k-I}^k)}\right) \\ & + \frac{((1-\rho_k) \hat{\sigma}^2(\hat{a}_{k-I}^k) \mathfrak{M}(S_{k-1}^k) + \rho_k \sigma^2(a_{k-I}^k) \hat{\mathfrak{M}}(\hat{S}_{k-1}^k))^2}{2\sigma^2(a_{k-I}^k) \hat{\sigma}^2(\hat{a}_{k-I}^k) ((1-\rho_k) \hat{\sigma}^2(\hat{a}_{k-I}^k) + \rho_k \sigma^2(a_{k-I}^k))}. \end{aligned}$$

REFERENCES

- [1] Kavcic Aleksandar and Jose M. F. Moura. The Viterbi Algorithm and Markov Noise Memory. *IEEE Trans. on Info. Th.*, 46, Issue: 1:291–301, 2000.
- [2] E. Biglieri. High-level modulation and coding for nonlinear satellite channels. *IEEE Transactions on Communications*, 32(5):616 – 626, may 1984.
- [3] Jr. Forney, G. Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference. *IEEE Trans. on Info. Th.*, 18, Issue: 3:363 – 378, 1972.
- [4] Jr. Forney, G.D. The viterbi algorithm. *Proceedings of the IEEE*, 61(3):268 – 278, march 1973.
- [5] R. G. Gallager. *Information Theory and Reliable Communication*. New York, NY: Wiley, 1968.
- [6] A. Kavcic and J.M.F. Moura. Correlation-sensitive adaptive sequence detection. *IEEE Transactions on Magnetics*, 34(3):763–771, may 1998.
- [7] N. Kumar, P. Agarwal, A. Ramamoorthy, and M. Salapaka. Maximum likelihood sequence detector for dynamic mode high density probe storage. *IEEE Transactions on Communication*, June 2010.
- [8] Y.-J. Liu, I. Oka, and E. Biglieri. Error probability for digital transmission over nonlinear channels with application to tcm. *IEEE Transactions on Information Theory*, 36(5):1101 –1110, sep 1990.
- [9] A. Viterbi. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, 13(2):260 – 269, apr 1967.
- [10] Andrew J. Viterbi and Jim K. Omura. *Principles Of Digital Communication And Coding*. McGraw-Hill, Inc. New York, NY, USA, 1979.
- [11] Jian-Gang Zhu and Haiyun Wang. Noise characteristics of interacting transitions in longitudinal thin film media. *IEEE Transactions on Magnetics*, 31(2):1065 –1070, mar 1995.